

TITLE OF THE INVENTION

PCA3 MESSENGER RNA SPECIES IN BENIGN AND MALIGNANT PROSTATE TISSUES

This application claims the benefit of the filing date of U.S. Provisional Application 60/156,549 filed September 29, 1999. The contents of this application is incorporated herein by reference.

5

FIELD OF THE INVENTION

10

15

20

The present invention relates to prostate cancer. More specifically, the present invention relates to nucleic acid molecules (messenger RNAs) encoded by the gene PCA3; the differential expression of two of these RNA species in non-malignant and malignant prostatic states, methods for specifically diagnosing prostate cancer based on the detection of the RNA species related to prostate cancer, therapeutic approaches to prostate cancer implying these two RNA species; nucleic acid molecules and antibodies having binding affinity for the differentially expressed mRNAs; kits containing nucleic acid probes or antibodies; bioassays using the nucleic acid sequences of the differentially expressed mRNAs of the present invention to diagnose, assess or prognose a mammal afflicted with or susceptible to developing prostate cancer, and to bioassays to screen for compounds which modulate the expression of the mRNAs of the present invention.

BACKGROUND OF THE INVENTION

25

Over the last decade, cancer of the prostate has become the most commonly diagnosed malignancy among men and the second leading cause of male cancer deaths in the western population, following lung cancer (Landis et al., 1998, CA Cancer J. Clin. 48(1):6-29). Of all cancers, the incidence of prostate cancer increases most rapidly with age. As longevity among the western population increases, there continues to be a corresponding rise in the number of prostate cancers with an expected

increase of 60% in this decade alone. Mortality has increased at a slower rate, but overall has doubled in the last 50 years. Although the disease is typically diagnosed in men over the age of 65, its impact is still significant in that the average life span of a man who dies from prostate cancer is reduced by 9-10 years. If discovered, early prostate cancer can now be cured with surgery in approximately 90% of cases. Unfortunately the disease is slowly fatal once the tumor spreads outside the area of the gland and forms distant metastases. Early detection of the disease, while still confined to the prostate gland, and accurate staging for the selection of appropriate therapy should improve mortality rates.

Despite many advances in recent years, the precision with which an individual suffering from prostate cancer can be staged is still sub-optimal. The main reason for this is that tumor spread beyond the prostate is generally microscopic rather than macroscopic. Digital rectal examination of the prostate has been the cornerstone for the local staging of prostatic cancer for many decades, but it oftentimes underestimates the extent of the disease. Transrectal ultrasound by itself is only of limited value as a means of prostate cancer staging. Computertomography and magnetic resonance imaging have generally been disappointing in the staging of prostate cancer (Kirby, 1997, Prostate cancer and Prostatic Diseases 1:2-10). Recent promising approaches to prostate cancer staging imply the use of biochemical and molecular technologies, centered around proteins or their corresponding nucleic acids which are preferentially expressed in prostate cells (Lange, 1997, In « Principles and Practice of Genitourinary Oncology » ed. Lippincott-Raven Publishers, Ch. 41, pp. 417-425). The most notorious prostate markers are PSA (prostate specific antigen) and PSM (prostate specific membrane) antigen

PSA is a secreted glycoprotein encoded by the PSA gene located on chromosome 19. It is expressed under androgen control by glandular epithelial cells of the prostate and secreted into seminal plasma. PSA protein is normally confined to the prostate but in the case of prostatic disease such as cancer or BPH (benign prostate hyperplasia), PSA leaks into the blood where it is present in different forms, including one that is and one that is not bound to protein complexes (El-Shirbiny, 1994, Adv. Clin. Chem. 31:99). The measurement of total PSA serum concentrations is one of the most frequently used and FDA-approved biochemical tests in the screening and management of prostate cancer patients. Studies to date have suggested that screening with PSA, in conjunction with digital rectal exams and transrectal ultrasound, increases the detection of early prostate cancers often while still localized to the gland itself (Brawer et al., 1992, J. Urol. 147:841). Serum PSA is also useful for monitoring of patients after therapy, especially after surgical prostatectomy. However, total PSA measurements also identify a large number of patients with abnormally elevated levels who are subsequently found to have no prostate cancer. Recently, the concept of measuring the percent free/total PSA ratio was shown to increase the specificity of prostate cancer screening in men with PSA between 4 and 10 ng/mL (Letran et al., 1998, J. Urol. 160:426).

The PSM gene encodes a transmembrane glycoprotein expressed by epithelial cells of normal prostate, benign prostate hyperplasia and, to a greater extent, malignant prostatic tissue. Low levels of PSM are also detected in some other tissues (Israeli et al., 1994, Cancer Res. 54:1807). PSA and PSM have also been targets for molecular approaches to prostate cancer using RT-PCR (reverse transcription-polymerase chain reaction). This very sensitive nucleic acid amplification technology is used to identify cells based on the expression of specific messenger RNAs. It

involves preparing RNA samples from tissues or body fluids, reverse transcribing it into cDNA and amplifying specific cDNAs by the use of primers that target the particular gene of interest. RT-PCR analyses of blood, lymph nodes and bone marrow from prostate cancer patients using PSA and PSM have disclosed the extreme sensitivity of this approach. However, the clinical value of molecular tests still has to be confirmed (Verkaik et al., 1997, Urol. Res. 25:373; Gomella et al., 1997, J. Urol. 158:326).

Thus, there remains a need to provide a more sensitive test for diagnosing prostate cancer. There also remains a need to provide a better test for the staging of prostate cancer. There also remains a need to provide a prostate cancer marker which is more specific and more reliable to prostate cancer detection, staging and treatment methods.

The present invention seeks to meet these and other needs.

A new prostate cancer marker, PCA3, was discovered a few years ago by differential display analysis intended to highlight genes associated with prostate cancer development (PCT application number PCT/CA98/00346). PCA3 is located on chromosome 9 and composed of four exons. It encodes at least four different transcripts which are generated by alternative splicing and polyadenylation. By RT-PCR analysis, PCA3 expression was found to be limited to the prostate and absent in all other tissues, including testis, ovary, breast and bladder. Northern blot analysis showed that PCA3 is highly expressed in the vast majority of prostate cancers examined (47 out of 50) whereas no or very low expression is detected in benign prostate hyperplasia or normal prostate cells from the same patients. There is at least 20-fold overexpression of PCA3 in prostatic carcinomas in comparison to normal or BPH tissues. PCA3 expression seems to increase with tumor grade and is detected in metastatic lesions

In summary, prostate cancer staging based on specific markers such as PSA and PSM is a very promising avenue for the management of the disease. The drawback of using PSA or PSM for prostate cancer staging is that they are expressed in normal as well as in cancerous cells. In addition, poorly differentiated tumors may escape diagnosis since they tend to produce significantly less PSA protein than less aggressive tumors. This is the case for 10% of all prostate cancers. PCA3, on the other hand, is differentially expressed in cancerous and normal prostate cells, and its expression does not decrease with increasing tumor grade. PCA3 could therefore be a useful tool which may overcome the drawbacks of PSA and PSM in the diagnosis, staging and treatment of prostate cancer patients.

The present description refers to a number of documents, the content of which is herein incorporated by reference, in their entirety.

SUMMARY OF THE INVENTION

The invention concerns the discovery of distinct PCA3 RNAs associated with a non-malignant and/or malignant status of the prostate.

The invention also concerns the identification that a balance between the level of these PCA3 mRNAs correlates with the non-malignant or the malignant status of the prostate.

One of these RNAs corresponds to a PCA3 RNA molecule having an additional sequence of 228 bp (shown in SEQ ID NO.1), inserted between exons 3 and 4a, whereas the other lacks the additional sequence (SEQ ID NO.2). The RNA lacking the additional sequence is associated with prostate cancer whereas the RNA comprising same is associated with a non-malignant prostatic state. Based on the differential expression of these two PCA3 RNA species, protocols for the diagnosis of prostate disease are

provided. The above findings could also lead to a therapeutic approach to prostate cancer.

The invention further concerns reagents and methods to assess the prostate status in an animal, comprising a quantitative determination of SEQ ID NO:1 or fragments, or variants thereof with respect to SEQ ID NO:2 or fragments, or variants thereof.

Thus, the present invention relates to the discovery and characterization of a novel sequence expressed in PCA3 mRNA, which enables a determination of the prostate status of an animal, based on a determination of the relative abundance of two differentially expressed PCA3 mRNAs.

The invention provides, in general, isolated nucleic acid molecules encoding differentially expressed PCA3 mRNAs and to variants or portions thereof, retaining their ability to enable a prostate status determination.

The invention further provides purified polypeptides encoded by the differentially expressed PCA3 mRNAs of the present invention or an epitope binding portion thereof.

The invention also provides nucleic acids for the specific detection of the presence of differentially expressed PCA3 mRNAs associated with prostate cancer or proteins or polypeptides encoded by such mRNAs in a sample.

The invention further provides a method of detecting nucleic acid encoding differentially expressed PCA3 mRNAs.

The invention also provides a kit for detecting the presence of nucleic acid encoding differentially-expressed PCA3 mRNAs in a sample.

The invention in addition provides a recombinant nucleic acid molecule comprising, 5' to 3', a promoter effective to initiate transcription in

a host cell and the above-described isolated nucleic acid molecule, variants or fragments thereof, encoding differentially expressed PCA3 mRNAs.

The invention also provides a recombinant nucleic acid molecule comprising a vector and the above-described isolated nucleic acid molecule encoding differentially expressed PCA3 mRNAs.

The invention further provides an antisense nucleic acid molecule specific to the differentially expressed PCA3 mRNAs.

The invention also provides a cell that contains an above-described recombinant nucleic acid molecule.

The invention further relates to a non-human organism that contains an above-described recombinant nucleic acid molecule encoding a differentially expressed PCA3 mRNA. In particular, the invention relates to a non-human organism containing a recombinant nucleic acid molecule encoding a PCA3 mRNA having an additional sequence between exon 3 and exon 4a. In a particularly preferred embodiment, this additional sequence comprises the sequence of SEQ ID NO 1, variants of parts thereof.

The invention also relates to an antibody having binding affinity specifically to a polypeptide encoded by a differentially expressed PCA3 mRNA or an epitope-bearing portion thereof.

The invention further provides a method of detecting differentially expressed PCA3 mRNAs in a sample. As well, it also provides a method of measuring the amount of differentially expressed PCA3 mRNAs in a sample.

The invention further relates to a method of detecting antibodies having binding affinity specifically to polypeptides encoded by a differentially expressed PCA3 mRNA.

In one embodiment, the invention further relates to a diagnostic kit comprising a first container means containing nucleic acid

molecules specific for a differentially expressed PCA3 mRNA, and a second container means containing a probe specific to a differentially expressed PCA3 mRNAs.

5 In another embodiment, the invention relates to a diagnostic kit comprising a first container means containing the above-described antibody, and a second container means containing a conjugate comprising a binding partner of the monoclonal antibody and a label.

The invention also provides a hybridoma which produces the above-described monoclonal antibody.

10 As well, the invention further relates to diagnostic methods for human disease, in particular, prostate cancer. Preferably, a method of diagnosing the presence or predisposition to develop prostate cancer in a patient is provided herein.

15 The invention also provides methods for therapeutic uses involving all or part of (1) a nucleic acid sequence encoding the differentially expressed PCA3 mRNAs, variants or parts thereof, (2) antisense to differentially expressed PCA3 mRNA molecules, variants or parts thereof, (3) protein encoded by a differentially expressed PCA3 mRNA, variants or parts thereof, or (4) antibodies to proteins encoded by differentially expressed PCA3 mRNAs.

20 Further, the invention provides a method to modulate the level of differentially expressed of a first PCA3 mRNA (e.g. the long one) and a second PCA3 mRNA (e.g. the short one) by expressing one of the first or second differentially expressed mRNA. In a preferred embodiment, the invention provides a modulation of the differentially expressed PCA3 mRNAs such that the level of the first PCA3 mRNA is superior to that of the second.

25 Having identified the differential expression of mRNAs as a marker for prostatic state of an animal, and more particularly having shown

that the presence of the additional sequence, which interrupts the coding sequence of the PCA3 encoded protein, correlates with a non-malignant state, while the absence of the additional sequence and a non-interruption of the thereby encoded protein, correlates with malignant cancer, the present invention therefore provides the means to interrupt the coding sequence of the PCA3 protein, using any means of genetic engineering, known to a skilled artisan, and assesses whether such an interruption can revert the malignant phenotype

In order to provide a clear and consistent understanding of terms used in the present description, a number of definitions are provided hereinbelow

As used herein, the terminology "non-malignant prostate or status" is meant to cover a non-cancerous prostatic state. Thus, these terminologies are meant to include a normal status as well as a benign prostatic status (such as BPH, for example)

Since the differentiating markers between the malignant and non-malignant prostatic state is at the mRNA and protein level (i.e. an expressed marker), one of the advantages of the present invention is to enable a determination of the prostatic status in an animal using a number of available means to the skilled artisan. Non-limiting examples of such means include nucleic acid probes, antibodies, ligands and PNAs, in easily obtainable cells which express these differentiating markers. A non-limiting example thereof is lymphocytes, thereby enabling a determination from a simple blood sample.

The term "sample" is used herein broadly to refer to all types of samples from an animal in which the differential expression of the short and/or long PCA3 nucleic acid or protein of the present invention can be

analyzed. Non-limiting examples thereof include biopsies, blood, fine needle aspirate, urine and bone marrow

Nucleotide sequences are presented herein by single strand, in the 5' to 3' direction, from left to right, using the one letter nucleotide symbols as commonly used in the art and in accordance with the recommendations of the IUPAC-IUB Biochemical Nomenclature Commission

Unless defined otherwise, the scientific and technological terms and nomenclature used herein have the same meaning as commonly understood by a person of ordinary skill to which this invention pertains. Generally, the procedures for cell cultures, infection, molecular biology methods and the like are common methods used in the art. Such standard techniques can be found in reference manuals such as for example Sambrook et al (1989, Molecular Cloning-A Laboratory Manual, Cold Spring Harbor Laboratories) and Ausubel et al (1994, Current Protocols in Molecular Biology, Wiley, New York).

The present description refers to a number of routinely used recombinant DNA (rDNA) technology terms. Nevertheless, definitions of selected examples of such rDNA terms are provided for clarity and consistency.

As used herein, "nucleic acid molecule", refers to a polymer of nucleotides. Non-limiting examples thereof include DNA (i.e. genomic DNA, cDNA) and RNA molecules (i.e. mRNA). The nucleic acid molecule can be obtained by cloning techniques or synthesized. DNA can be double-stranded or single-stranded (coding strand or non-coding strand [antisense])

The term "recombinant DNA" as known in the art refers to a DNA molecule resulting from the joining of DNA segments. This is often referred to as genetic engineering.

5 The term "DNA segment", is used herein, to refer to a DNA molecule comprising a linear stretch or sequence of nucleotides. This sequence when read in accordance with the genetic code, can encode a linear stretch or sequence of amino acids which can be referred to as a polypeptide, protein, protein fragment and the like.

10 The terminology "amplification pair" refers herein to a pair of oligonucleotides (oligos) of the present invention, which are selected to be used together in amplifying a selected nucleic acid sequence by one of a number of types of amplification processes, preferably a polymerase chain reaction. Other types of amplification processes include ligase chain reaction, strand displacement amplification, or nucleic acid sequence-based amplification, as explained in greater detail below. As commonly known in the art, the oligos are designed to bind to a complementary sequence under selected conditions

15
20 In one particular embodiment, amplification of a nucleic acid sample from a patient is amplified under conditions which favor the amplification of the most abundant differentially expressed nucleic acid. In one preferred embodiment, RT-PCR is carried out on an mRNA sample from a patient under conditions which favor the amplification of the most abundant PCA3 mRNA. In another preferred embodiment, the amplification of the differentially expressed PCA3 nucleic acids is carried out simultaneously. Of
25 course, it will be realized by a person skilled in the art that such methods could be adapted for the detection of differentially expressed proteins instead of differentially expressed nucleic acid sequences

The nucleic acid (i.e. DNA or RNA) for practicing the present invention may be obtained according to well known methods

Oligonucleotide probes or primers of the present invention may be of any suitable length, depending on the particular assay format and the particular needs and targeted genomes employed. In general, the oligonucleotide probes or primers are at least 12 nucleotides in length, preferably between 15 and 24 molecules, and they may be adapted to be especially suited to a chosen nucleic acid amplification system. As commonly known in the art, the oligonucleotide probes and primers can be designed by taking into consideration the melting point of hybridization thereof with its targeted sequence (see below and in Sambrook et al., 1989, Molecular Cloning - A Laboratory Manual, 2nd Edition, CSH Laboratories; Ausubel et al., 1989, in Current Protocols in Molecular Biology, John Wiley & Sons Inc., N.Y.).

The term "oligonucleotide" or "DNA" molecule or sequence refers to a molecule comprised of the deoxyribonucleotides adenine (A), guanine (G), thymine (T) and/or cytosine (C). When in a double-stranded form, it can comprise or include a "regulatory element" according to the present invention, as the term is defined herein. The term "oligonucleotide" or "DNA" can be found in linear DNA molecules or fragments, viruses, plasmids, vectors, chromosomes or synthetically derived DNA. As used herein, particular double-stranded DNA sequences may be described according to the normal convention of giving only the sequence in the 5' to 3' direction. It will also be recognized that "oligonucleotide" can be in a single-stranded form.

"Nucleic acid hybridization" refers generally to the hybridization of two single-stranded nucleic acid molecules having complementary base sequences, which under appropriate conditions will

form a thermodynamically favored double-stranded structure. Examples of hybridization conditions can be found in the two laboratory manuals referred above (Sambrook et al., 1989, *supra* and Ausubel et al., 1989, *supra*) and are commonly known in the art. In the case of a hybridization to a nitrocellulose filter, as for example in the well known Southern blotting procedure, a nitrocellulose filter can be incubated overnight at 65°C with a labeled probe in a solution containing 50% formamide, high salt (5 x SSC or 5 x SSPE), 5 x Denhardt's solution, 1% SDS, and 100 µg/ml denatured carrier DNA (i.e. salmon sperm DNA). The non-specifically binding probe can then be washed off the filter by several washes in 0.2 x SSC/0.1% SDS at a temperature which is selected in view of the desired stringency: room temperature (low stringency), 42°C (moderate stringency) or 65°C (high stringency). The selected temperature is based on the melting temperature (T_m) of the DNA hybrid. Of course, RNA-DNA hybrids can also be formed and detected. In such cases, the conditions of hybridization and washing can be adapted according to well-known methods by the person of ordinary skill. Stringent conditions will be preferably used (Sambrook et al., 1989, *supra*).

Probes of the invention can be utilized with naturally occurring sugar-phosphate backbones as well as modified backbones including phosphorothioates, dithionates, alkyl phosphonates and α -nucleotides and the like. Modified sugar-phosphate backbones are generally taught by Miller, 1988, Ann. Reports Med. Chem. 23:295 and Moran et al., 1987, Nucleic acid molecule Acids Res., 14:5019. Probes of the invention can be constructed of either ribonucleic acid (RNA) or deoxyribonucleic acid (DNA), and preferably of DNA.

The types of detection methods in which probes can be used include Southern blots (DNA detection), dot or slot blots (DNA, RNA), and Northern blots (RNA detection). Although less preferred, labeled proteins

could also be used to detect a particular nucleic acid sequence to which it binds. More recently, PNAs have been described (Nielsen et al. 1999, Current Opin. Biotechnol. 10:71-75). PNAs could also be used to detect the mRNAs of the present invention. Other detection methods include kits containing probes on a dipstick setup and the like

Although the present invention is not specifically dependent on the use of a label for the detection of a particular nucleic acid sequence, such a label might be beneficial, by increasing the sensitivity of the detection. Furthermore, it enables automation. Probes can be labeled according to numerous well known methods (Sambrook et al., 1989, *supra*). Non-limiting examples of labels include ^3H , ^{14}C , ^{32}P , and ^{35}S . Non-limiting examples of detectable markers include ligands, fluorophores, chemiluminescent agents, enzymes, and antibodies. Other detectable markers for use with probes, which can enable an increase in sensitivity of the method of the invention, include biotin and radionucleotides. It will become evident to the person of ordinary skill that the choice of a particular label dictates the manner in which it is bound to the probe.

As commonly known, radioactive nucleotides can be incorporated into probes of the invention by several methods. Non-limiting examples thereof include kinasing the 5' ends of the probes using gamma ^{32}P ATP and polynucleotide kinase, using the Klenow fragment of Pol I of *E. coli* in the presence of radioactive dNTP (i.e. uniformly labeled DNA probe using random oligonucleotide primers in low-melt gels), using the SP6/T7 system to transcribe a DNA segment in the presence of one or more radioactive NTP, and the like

As used herein, "oligonucleotides" or "oligos" define a molecule having two or more nucleotides (ribo or deoxyribonucleotides). The size of the oligo will be dictated by the particular situation and ultimately on

the particular use thereof and adapted accordingly by the person of ordinary skill. An oligonucleotide can be synthesised chemically or derived by cloning according to well known methods

As used herein, a "primer" defines an oligonucleotide which is capable of annealing to a target sequence, thereby creating a double stranded region which can serve as an initiation point for DNA synthesis under suitable conditions.

Amplification of a selected, or target, nucleic acid sequence may be carried out by a number of suitable methods. See generally Kwok et al., 1990, Am. Biotechnol. Lab 8:14-25. Numerous amplification techniques have been described and can be readily adapted to suit particular needs of a person of ordinary skill. Non-limiting examples of amplification techniques include polymerase chain reaction (PCR), ligase chain reaction (LCR), strand displacement amplification (SDA), transcription-based amplification, the Q β replicase system and NASBA (Kwok et al., 1989, Proc. Natl. Acad. Sci. USA 86, 1173-1177; Lizardi et al., 1988, BioTechnology 6:1197-1202; Malek et al., 1994, Methods Mol. Biol. 28:253-260; and Sambrook et al., 1989, *supra*). Preferably, amplification will be carried out using PCR.

Polymerase chain reaction (PCR) is carried out in accordance with known techniques. See, e.g., U.S. Pat. Nos. 4,683,195, 4,683,202, 4,800,159; and 4,965,188 (the disclosures of all three U.S. Patent are incorporated herein by reference). In general, PCR involves, a treatment of a nucleic acid sample (e.g., in the presence of a heat stable DNA polymerase) under hybridizing conditions, with one oligonucleotide primer for each strand of the specific sequence to be detected. An extension product of each primer which is synthesized is complementary to each of the two nucleic acid strands, with the primers sufficiently complementary to each strand of the specific sequence to hybridize therewith. The extension product

synthesized from each primer can also serve as a template for further synthesis of extension products using the same primers. Following a sufficient number of rounds of synthesis of extension products, the sample is analysed to assess whether the sequence or sequences to be detected are present. Detection of the amplified sequence may be carried out by visualization following EtBr staining of the DNA following gel electrophores, or using a detectable label in accordance with known techniques, and the like. For a review on PCR techniques (see PCR Protocols, A Guide to Methods and Amplifications, Michael et al. Eds, Acad. Press, 1990).

Ligase chain reaction (LCR) is carried out in accordance with known techniques (Weiss, 1991, Science 254:1292). Adaptation of the protocol to meet the desired needs can be carried out by a person of ordinary skill. Strand displacement amplification (SDA) is also carried out in accordance with known techniques or adaptations thereof to meet the particular needs (Walker et al., 1992, Proc. Natl. Acad. Sci. USA 89:392-396, and *ibid.*, 1992, Nucleic Acids Res. 20:1691-1696).

As used herein, the term "gene" is well known in the art and relates to a nucleic acid sequence defining a single protein or polypeptide. A "structural gene" defines a DNA sequence which is transcribed into RNA and translated into a protein having a specific amino acid sequence thereby giving rise to a specific polypeptide or protein. It will be readily recognized by the person of ordinary skill, that the nucleic acid sequence of the present invention can be incorporated into anyone of numerous established kit formats which are well known in the art.

A "heterologous" (i.e. a heterologous gene) region of a DNA molecule is a subsegment segment of DNA within a larger segment that is not found in association therewith in nature. The term "heterologous" can be similarly used to define two polypeptidic segments not joined together in

nature. Non-limiting examples of heterologous genes include reporter genes such as luciferase, chloramphenicol acetyl transferase, β -galactosidase, and the like which can be juxtaposed or joined to heterologous control regions or to heterologous polypeptides.

5 The term "vector" is commonly known in the art and defines a plasmid DNA, phage DNA, viral DNA and the like, which can serve as a DNA vehicle into which DNA of the present invention can be cloned. Numerous types of vectors exist and are well known in the art.

10 The term "expression" defines the process by which a gene is transcribed into mRNA (transcription), the mRNA is then being translated (translation) into one polypeptide (or protein) or more.

15 The terminology "expression vector" defines a vector or vehicle as described above but designed to enable the expression of an inserted sequence following transformation into a host. The cloned gene (inserted sequence) is usually placed under the control of control element sequences such as promoter sequences. The placing of a cloned gene under such control sequences is often referred to as being operably linked to control elements or sequences.

20 Operably linked sequences may also include two segments that are transcribed onto the same RNA transcript. Thus, two sequences, such as a promoter and a "reporter sequence" are operably linked if transcription commencing in the promoter will produce an RNA transcript of the reporter sequence. In order to be "operably linked" it is not necessary that two sequences be immediately adjacent to one another.

25 Expression control sequences will vary depending on whether the vector is designed to express the operably linked gene in a prokaryotic or eukaryotic host or both (shuttle vectors) and can additionally contain transcriptional elements such as enhancer elements, termination sequences,

tissue-specificity elements, and/or translational initiation and termination sites.

Prokaryotic expressions are useful for the preparation of large quantities of the protein encoded by the DNA sequence of interest. This protein can be purified according to standard protocols that take advantage of the intrinsic properties thereof, such as size and charge (i.e. SDS gel electrophoresis, gel filtration, centrifugation, ion exchange chromatography...) In addition, the protein of interest can be purified via affinity chromatography using polyclonal or monoclonal antibodies. The purified protein can be used for therapeutic applications

The DNA construct can be a vector comprising a promoter that is operably linked to an oligonucleotide sequence of the present invention, which is in turn, operably linked to a heterologous gene, such as the gene for the luciferase reporter molecule. "Promoter" refers to a DNA regulatory region capable of binding directly or indirectly to RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of the present invention, the promoter is bound at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter will be found a transcription initiation site (conveniently defined by mapping with S1 nuclease), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CCAT" boxes. Prokaryotic promoters contain -10 and -35 consensus sequences, which serve to initiate transcription and the transcript products contain a Shine-Dalgarno sequence, which serves as a ribosome binding sequence during translation initiation

As used herein, the designation "functional derivative" denotes, in the context of a functional derivative of a sequence whether an nucleic acid or amino acid sequence, a molecule that retains a biological activity (either function or structural) that is substantially similar to that of the original sequence. This functional derivative or equivalent may be a natural derivative or may be prepared synthetically. Such derivatives include amino acid sequences having substitutions, deletions, or additions of one or more amino acids, provided that the biological activity of the protein is conserved. The same applies to derivatives of nucleic acid sequences which can have substitutions, deletions, or additions of one or more nucleotides, provided that the biological activity of the sequence is generally maintained. When relating to a protein sequence, the substituting amino acid as chemico-physical properties which are similar to that of the substituted amino acid. The similar chemico-physical properties include, similarities in charge, bulkiness, hydrophobicity, hydrophylicity and the like. The term "functional derivatives" is intended to include "fragments", "segments", "variants", "analogs" or "chemical derivatives" of the subject matter of the present invention.

Thus, the term "variant" refers herein to a protein or nucleic acid molecule which is substantially similar in structure and biological activity to the protein or nucleic acid of the present invention.

The functional derivatives of the present invention can be synthesized chemically or produced through recombinant DNA technology. All these methods are well known in the art.

As used herein, "chemical derivatives" is meant to cover additional chemical moieties not normally part of the subject matter of the invention. Such moieties could affect the physico-chemical characteristic of the derivative (i.e. solubility, absorption, half life and the like, decrease of

toxicity) Such moieties are exemplified in Remington's Pharmaceutical Sciences (1980) Methods of coupling these chemical-physical moieties to a polypeptide are well known in the art.

5 The term "allele" defines an alternative form of a gene which occupies a given locus on a chromosome

As commonly known, a "mutation" is a detectable change in the genetic material which can be transmitted to a daughter cell. As well known, a mutation can be, for example, a detectable change in one or more deoxyribonucleotide. For example, nucleotides can be added, deleted, substituted for, inverted, or transposed to a new position. Spontaneous mutations and experimentally induced mutations exist. The result of a mutations of nucleic acid molecule is a mutant nucleic acid molecule. A mutant polypeptide can be encoded from this mutant nucleic acid molecule

10

As used herein, the term "purified" refers to a molecule having been separated from a cellular component. Thus, for example, a "purified protein" has been purified to a level not found in nature. A "substantially pure" molecule is a molecule that is lacking in all other cellular components

15

As used herein, the terms "molecule", "compound", or "agent" are used interchangeably and broadly to refer to natural, synthetic or semi-synthetic molecules or compounds. The term "molecule" therefore denotes for example chemicals, macromolecules, cell or tissue extracts (from plants or animals) and the like. Non limiting examples of molecules include nucleic acid molecules, peptides, ligands, including antibodies, carbohydrates and pharmaceutical agents. The agents can be selected and screened by a variety of means including random screening, rational selection and by rational design using for example protein or ligand modelling methods such as computer modelling. The terms "rationally selected" or

20

25

"rationally designed" are meant to define compounds which have been chosen based on the configuration of the interaction domains of the present invention. As will be understood by the person of ordinary skill, macromolecules having non-naturally occurring modifications are also within the scope of the term "molecule". For example, peptidomimetics, well known in the pharmaceutical industry and generally referred to as peptideanalogs can be generated by modelling as mentioned above. Similarly, in a preferred embodiment, the polypeptides of the present invention are modified to enhance their stability. It should be understood that in most cases this modification should not alter the biological activity of the protein. The molecules identified in accordance with the teachings of the present invention have a therapeutic value in diseases or conditions in which the physiology or homeostasis of the cell and/or tissue is compromised by a defect in the expression of PCA3 mRNAs. Alternatively, the molecules identified in accordance with the teachings of the present invention find utility in the development of compounds which can modulate the expression of a differentially expressed PCA3 mRNA or modulate the activity or level of a protein encoded thereby.

As used herein, agonists and antagonists also include potentiators of known compounds with such agonist or antagonist properties. In one embodiment, modulators of the level or the activity of the PCA3 protein lacking the additional sequence of the present invention can be identified and selected by contacting the indicator cell with a compound or mixture or library of molecules for a fixed period of time. In certain embodiments, the additional sequence-containing PCA3 protein can serve as a control.

The present invention also provides antisense nucleic acid molecules which can be used for example to decrease or abrogate the

expression of the PCA3 mRNA lacking the additional sequence of the present invention or of the protein encoded thereby. An antisense nucleic acid molecule according to the present invention refers to a molecule capable of forming a stable duplex or triplex with a portion of its targeted nucleic acid sequence (DNA or RNA). The use of antisense nucleic acid molecules and the design and modification of such molecules is well known in the art as described for example in WO 96/32966, WO 96/11266, WO 94/15646, WO 93/08845 and USP 5,593,974. Antisense nucleic acid molecules according to the present invention can be derived from the nucleic acid sequences and modified in accordance to well known methods. For example, some antisense molecules can be designed to be more resistant to degradation to increase their affinity to their targeted sequence, to affect their transport to chosen cell types or cell compartments, and/or to enhance their lipid solubility by using nucleotide analogs and/or substituting chosen chemical fragments thereof, as commonly known in the art.

An indicator cell in accordance with the present invention can be used to identify antagonists. For example, the test molecule or molecules are incubated with the host cell in conjunction with one or more agonists held at a fixed concentration. An indication and relative strength of the antagonistic properties of the molecule(s) can be provided by comparing the level of gene expression in the indicator cell in the presence of the agonist, in the absence of test molecules vs in the presence thereof. Of course, the antagonistic effect of a molecule can also be determined in the absence of agonist, simply by comparing the level of expression of the reporter gene product in the presence and absence of the test molecule(s).

It shall be understood that the *'in vivo'* experimental model can also be used to carry out an *'in vitro'* assay. For example, cellular

extracts from the indicator cells can be prepared and used in one of the aforementioned "in vitro" tests

As used herein the recitation "indicator cells" refers to cells that express a differentially expressed PCA3 mRNA according to the present invention. In some embodiment, the protein encoded by the nucleic acid sequence can be coupled to an identifiable or selectable phenotype or characteristic. Such indicator cells can be used in the screening assays of the present invention. In certain embodiments, the indicator cells have been engineered so as to express a chosen derivative, fragment, homolog, or mutant of the differentially-expressed PCA3 mRNA of the present invention. The cells can be yeast cells or higher eukaryotic cells such as mammalian cells. When the binding partner for the PCA3 protein will have been identified, the interaction between the two partners will be able to serve as a target for the modulation of the activity of this PCA3-encoded protein. In one particular embodiment, the indicator cell would be a yeast cell harboring vectors enabling the use of the two-hybrid system technology, as well known in the art (Ausubel et al., 1994, *supra*) and can be used to test a compound or a library thereof. In one embodiment, a reporter gene encoding a selectable marker or an assayable protein can be operably linked to a control element such that expression of the selectable marker or assayable protein is dependent on the interaction of the PCA3-encoded protein and its binding partner. Such an indicator cell could be used to rapidly screen at high-throughput a vast array of test molecules. In a particular embodiment, the reporter gene is luciferase or β -Gal.

In some embodiments, it might be beneficial to express a protein of the present invention as a fusion protein. The design of constructs therefor and the expression and production of fusion proteins are well

known in the art (Sambrook et al., 1989, *supra*; and Ausubel et al., 1994, *supra*)

Non limiting examples of such fusion proteins include a hemagglutinin fusions and Gluthione-S-transferase (GST) fusions and Matrose binding protein (MBP) fusions. In certain embodiments, it might be beneficial to introduce a protease cleavage site between the two polypeptide sequences which have been fused. Such protease cleavage sites between two heterologously fused polypeptides are well known in the art.

In certain embodiments, it might also be beneficial to fuse the protein of the present invention to signal peptide sequences enabling a secretion of the fusion protein from the host cell. Signal peptides from diverse organisms are well known in the art. Bacterial OmpA and yeast Suc2 are two non limiting examples of proteins containing signal sequences. In certain embodiments, it might also be beneficial to introduce a linker (commonly known) between the interaction domain and the heterologous polypeptide portion. Such fusion protein find utility in the assays of the present invention as well as for purification purposes, detection purposes and the like.

For certainty, the sequences and polypeptides useful to practice the invention include without being limited thereto mutants, homologs, subtypes, alleles and the like. It shall be understood that generally, the sequences of the present invention should encode a functional (albeit defective) PCA3 protein. It will be clear to the person of ordinary skill that whether the PCA3 of the present invention, variant, derivative, or fragment thereof retains its can be determined by using the teachings and assays of the present invention and the general teachings of the art.

As exemplified herein below, the PCA3 protein of the present invention can be modified, for example by *in vitro* mutagenesis, to dissect the

structure-function relationship thereof and permit a better design and identification of modulating compounds. However, some derivative or analogs having lost their biological function may still find utility, for example for raising antibodies. These antibodies could be used for detection or purification purposes. In addition, these antibodies could also act as competitive or non-competitive inhibitor and be found to be modulators of the activity of the PCA3 protein of the present invention.

A host cell or indicator cell has been "transfected" by exogenous or heterologous DNA (e.g. a DNA construct) when such DNA has been introduced inside the cell. The transfecting DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes, yeast, and mammalian cells for example, the transfecting DNA may be maintained on a episomal element such as a plasmid. With respect to eukaryotic cells, a stably transfected cell is one in which the transfecting DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transfecting DNA. Transfection methods are well known in the art (Sambrook et al , 1989, *supra*, Ausubel et al., 1994 *supra*). The use of a mammalian cell as indicator can provide the advantage of furnishing an intermediate factor, which permits for example the interaction of two polypeptides which are tested, that might not be present in lowereukaryotes or prokaryotes. It will be understood that extracts from mammalian cells for example could be used in certain embodiments, to compensate for the lack of certain factors.

In general, techniques for preparing antibodies (including monoclonal antibodies and hybridomas) and for detecting antigens using

antibodies are well known in the art (Campbell, 1984, In "Monoclonal Antibody Technology. Laboratory Techniques in Biochemistry and Molecular Biology", Elsevier Science Publisher, Amsterdam, The Netherlands) and in Harlow et al., 1988 (in: Antibody-A Laboratory Manual, CSH Laboratories).
5 The present invention also provides polyclonal, monoclonal antibodies, or humanized versions thereof, chimeric antibodies and the like which inhibit or neutralize their respective interaction domains and/or are specific thereto.

From the specification and appended claims, the term therapeutic agent should be taken in a broad sense so as to also include a
10 combination of at least two such therapeutic agents. Further, the DNA segments or proteins according to the present invention can be introduced into individuals in a number of ways. For example, prostatic cells can be isolated from the afflicted individual, transformed with a DNA construct according to the invention and reintroduced to the afflicted individual in a
15 number of ways. Alternatively, the DNA construct can be administered directly to the afflicted individual. The DNA construct can also be delivered through a vehicle such as a liposome, which can be designed to be targeted to a specific cell type, and engineered to be administered through different routes.

20 For administration to humans, the prescribing medical professional will ultimately determine the appropriate form and dosage for a given patient, and this can be expected to vary according to the chosen therapeutic regimen (i.e. DNA construct, protein, cells), the response and condition of the patient as well as the severity of the disease.

25 Composition within the scope of the present invention should contain the active agent (i.e. fusion protein, nucleic acid, and molecule) in an amount effective to achieve the desired therapeutic effect while avoiding adverse side effects. Typically, the nucleic acids in accordance with the

present invention can be administered to mammals (i.e. humans) in doses ranging from 0.005 to 1 mg per kg of body weight per day of the mammal which is treated. Pharmaceutically acceptable preparations and salts of the active agent are within the scope of the present invention and are well known in the art (Remington's Pharmaceutical Science, 16th Ed., Mack Ed.). For the administration of polypeptides, antagonists, agonists and the like, the amount administered should be chosen so as to avoid adverse side effects. The dosage will be adapted by the clinician in accordance with conventional factors such as the extent of the disease and different parameters from the patient. Typically, 0.001 to 50 mg/kg/day will be administered to the mammal.

The present invention relates to a kit for diagnosing and/or staging prostate cancer or a predisposition to contracting same comprising a nucleic acid, a protein or a ligand in accordance with the present invention. For example, a compartmentalized kit in accordance with the present invention includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample (DNA, protein or cells), a container which contains the primers used in the assay, containers which contain enzymes, containers which contain wash reagents, and containers which contain the reagents used to detect the extension products.

The present invention also relates to a kit comprising the oligonucleotide primer of the present invention, which are specific to either

one of the PCA3 mRNA lacking the additional sequence of the present invention or the PCA3 mRNA containing the additional sequence of the present invention

Further objects and advantages of the present invention will be clear from the description that follows.

BRIEF DESCRIPTION OF THE DRAWINGS

Having thus generally described the invention, reference will now be made to the accompanying drawings, showing by way of illustration a preferred embodiment thereof, and in which.

Figure 1 shows the PCA3 genomic structure and location of oligonucleotides used for PCR;

Figure 2 shows a gel separating PCA3 RT-PCR products amplified from tissue biopsies of prostate cancer and benign prostate hyperplasia, using the primers of Example 1;

Figure 3 illustrates the nucleic acid sequences of RT-PCR-amplified PCA3 fragments with and without the additional sequence of the present invention. The sequences were amplified using PCR primers located in exon 3 and exon 4a. Primer sequences are shown in bold letters. Capital letters represent nucleic acids common to both sequences;

Figure 4 shows the amino acid sequence predicted from the PCA3 mRNAs containing the additional sequence of the present invention. This sequence corresponds to amino acids 1-23 of the original PCA3 polypeptide; and

Figure 5 shows examples of antigenic epitope-bearing PCA3 peptides comprising 8 amino acids (calculated according to H.G. Rammensee at <http://134.2.96.221/scripts/hlaserver.dll/EpPredict.htm>).

Other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of preferred embodiments with reference to the accompanying drawing which is exemplary and should not be interpreted as limiting the scope of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENT

In one embodiment, the present invention relates to an isolated and/or purified differentially-expressed PCA3 mRNA molecule. Preferably, the PCA3 mRNA or nucleic acid molecule comprises a polynucleotide sequence at least 90% identical (more preferably, 95%, 96%, 97%, 98%, 99% or 100% identical) to a sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a differentially expressed PCA3 polypeptide comprising the complete amino acid sequence in SEQ ID NO: 3;

(b) a nucleotide sequence complementary to any of the nucleotide sequences in (a) or (b).

In one preferred embodiment, the isolated nucleic acid molecule comprises a differentially expressed PCA3 mRNA nucleotide sequence with greater than 90% identity or similarity to the nucleotide sequence present in SEQ ID NO: 1 (preferably greater than 95%, 96%, 97%, 98%, 99% or 100%). In another preferred embodiment, the isolated nucleic acid molecule comprises the differentially expressed PCA3 mRNA sequence lacking the additional sequence present in SEQ ID NO:2. In another embodiment, the isolated differentially expressed additional sequence-minus mRNA sequence nucleic acid molecule encodes the differentially expressed PCA3 amino acid sequence present in SEQ ID NO:3.

Although PCT application CA98/00346 teaches a number of alternatively spliced mRNAs, prior to the present invention, a PCA3 mRNA comprising an additional sequence between exon 3 and exon 4a had not been identified. Furthermore, the identification of this additional sequence as a distinguishing marker of the prostate state had not been made. In addition, the correlation between the PCA3 mRNA minus the additional sequence and prostate cancer (as opposed to the PCA3 mRNA-containing the additional sequence in non-prostate cancer [i.e. normal or BPH]) had not been made. Thus, the additional sequence in PCA3 mRNA enables a prognosis and diagnosis of prostatic diseases in a patient. Preferably, the PCA3 nucleic acid molecule comprises a polynucleotide sequence at least 90% identical (more preferably, 95%, 96%, 97%, 98%, 99% or 100% identical) to one of the above-described differentially expressed mRNAs.

Also included within the scope of this invention are the functional equivalents of the herein-described isolated nucleic acid molecules and derivatives thereof. For example, the nucleic acid sequences depicted in SEQ ID NO:1 and SEQ ID NO:2 can be altered by substitutions, additions or deletions that provide for functionally equivalent molecules. Due to the degeneracy of nucleotide coding sequences, other DNA sequences which encode substantially the same amino acid sequence as depicted in SEQ ID NO: 3 can be used in the practice of the present invention.

In addition, the nucleic acid sequence can comprise a nucleotide sequence which results from the addition, deletion or substitution of at least one nucleotide to the 5'-end and/or the 3'-end of the nucleic acid formula shown in SEQ ID NO:1 or 2 or a derivative thereof. Any nucleotide or polynucleotide can be used in this regard, provided that its addition, deletion or substitution does not substantially alter the amino acid sequence of SEQ ID NO:3 which is encoded by the additional sequence-containing

nucleotide sequence. Moreover, the nucleic acid molecule of the present invention can, as necessary, have restriction endonuclease recognition sites added to its 5'-end and/or 3'-end. All variations of the nucleotide sequence of the PCA3 nucleotide coding sequence and fragments thereof permitted by the genetic code are, therefore, included in this invention.

Further, it is possible to delete codons or to substitute one or more codons by codons other than degenerate codons to produce a structurally modified polypeptide, but one which has substantially the same utility or activity of the polypeptide produced by the unmodified nucleic acid molecule. As recognized in the art, the two polypeptides are functionally equivalent, as are the two nucleic acid molecules which give rise to their production, even though the differences between the nucleic acid molecules are not related to degeneracy of the genetic code

One skilled in the art will realize that genomes often contain slight allelic variations between individuals. Therefore, the isolated nucleic acid molecule is also intended to include allelic variations, so long as the sequence is a functional derivative of the differentially expressed PCA3 mRNA coding sequence. When a PCA3 allele does not encode the identical sequence to that found in SEQ ID Nos 1 or 2, it can be isolated and identified as PCA3 using the same techniques used herein, and especially PCR techniques to amplify the appropriate gene with primers based on the sequences disclosed herein

One skilled in the art will realize that organisms other than humans might also contain differentially-expressed PCA3 mRNAs (for example, eukaryotes; more specifically, mammals, birds, fish, and plants; more specifically, gorillas, rhesus monkeys, and chimpanzees) The invention is intended to include, but not be limited to, differentially-expressed PCA3 mRNAs isolated from the above-described organisms

Isolated nucleic acid molecules of the present invention are also meant to include those chemically synthesized. For example, a nucleic acid molecule with the nucleotide sequence described herein or encoding the herein described differentially expressed products of PCA3 gene can be designed and, if necessary, divided into appropriate smaller fragments. Then an oligomer which corresponds to the nucleic acid molecule, or to each of the divided fragments, can be synthesized. Such syntheticoligonucleotides can be prepared, for example, by the triester method of Matteucci *et al.*, *J. Am. Chem. Soc.* 103:3185-3191 (1981) or by using an automated DNA synthesizer.

In another embodiment, the present invention relates to purified differentially expressed polypeptides (preferably, substantially pure) having an amino acid sequence corresponding to the herein described PCA3, or a functional derivative thereof. In a preferred embodiment, the polypeptide has the amino acid sequence set forth in SEQ ID NO: 3 or mutant or species variation thereof, or at least 80% identity or at least 90% similarity thereof (preferably, at least 90%, 95%, 96%, 97%, 98%, or 99% identity or at least 95%, 96%, 97%, 98%, or 99% similarity thereof), or at least 6 contiguous amino acids thereof (preferably, at least 10, 15, 20, 25, or 50 contiguous amino acids thereof).

In a preferred embodiment, the invention relates to differentially-expressed PCA3 epitopes. The epitope of these polypeptides is an immunogenic or antigenic epitope. An immunogenic epitope is that part of the protein which elicits an antibody response when the whole protein is the immunogen. An antigenic epitope is a fragment of the protein which can elicit an antibody response. Methods of selecting antigenic epitope fragments are well known in the art. See, Sutcliffe *et al.*, *Science* 219:660-666 (1983). Antigenic epitope-bearing peptides and polypeptides of the

invention are useful to raise an immune response that specifically recognizes the polypeptides. Antigenic epitope-bearing peptides and polypeptides of the invention comprise at least 7 amino acids (preferably, 9, 10, 12, 15 or 20 amino acids) of the proteins of the invention. Examples of an antigenic peptide are shown in Figure 5 as predicted using the method of Rammensee at <http://134.2.96.221/scripts/htaserver.dIREpPredict.htm>. Of course, it will be realized that other epitope-bearing PCA3 peptides could be predicted and used to raise antibodies.

A variety of methodologies known in the art can be utilized to obtain the peptide of the present invention. In one embodiment, the peptide is purified from tissues or cells which naturally produce the peptide. Alternatively, the above-described isolated nucleic acid fragments can be used to express the differentially expressed PCA3 protein in any organism. The samples of the present invention include cells, protein extracts or membrane extracts of cells, or biological fluids. The sample will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts used as the sample.

Any organism can be used as a source for the peptide of the invention, as long as the source organism naturally contains such a peptide. As used herein, "source organism" refers to the original organism from which the amino acid sequence of the subunit is derived, regardless of the organism the subunit is expressed in and ultimately isolated from.

In another embodiment, the present invention relates to a nucleic acid for the specific detection of the presence of PCA3 nucleic acid in a sample comprising the above-described nucleic acid molecules or at least a fragment thereof which binds under stringent conditions to PCA3 nucleic acid.

In one preferred embodiment, the present invention relates to nucleic acid probes which are complementary to a nucleotide sequence consisting of at least 10 consecutive nucleotides (preferably, 15, 18, 20, 25, or 30) from the nucleic acid molecule comprising a polynucleotide sequence at least 90% identical to a sequence selected from the group consisting of

(a) a nucleotide sequence encoding the PCA3 polypeptide comprising the complete amino acid sequence in SEQ ID NO. 3,

(b) a nucleotide sequence encoding the PCA3 gene comprising the nucleotide sequence in SEQ ID NO: 1, or 2;

(c) a nucleotide sequence complementary to any of the nucleotide sequences in (a) or (b), and

(d) a nucleotide sequence as previously described above.

In one embodiment of the above described method, a nucleic acid probe is immobilized on a solid support. Examples of such solid supports include, but are not limited to, plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, and acrylic resins, such as polyacrylamide and latex beads. Techniques for coupling nucleic acid probes to such solid supports are well known in the art.

The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The sample used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample which is compatible with the method utilized.

In another embodiment, the present invention relates to a method of detecting the presence of differentially expressed PCA3 mRNA in a sample comprising : a) contacting the sample with the above-described

nucleic acid probe, under specific hybridization conditions such that hybridization occurs, and b) detecting the presence of the probe bound to the nucleic acid molecule. One skilled in the art would select the nucleic acid probe according to techniques known in the art as described above.

5 Samples to be tested include but should not be limited to RNA samples from human tissue

10 Having identified that the additional PCA3 sequence of the present invention can be used as a marker for distinguishing between malignant and non malignant prostate states, probes which are specific to this additional sequence (or variants or fragments thereof) could also be used in accordance with the present invention. Of course, since in certain embodiments, such probes might detect genomic DNA, a positive signal coming from the genomic DNA might have to be eliminated in order to specifically detect the differentially expressed PCA3 mRNA

15 Although the present invention is specifically demonstrated using primers hybridizing to exon 3 and exon 4a sequences, it should be clear to the skilled artisan that primers derived from other regions of PCA3 could be used. For example, primers could be derived from sequences of exon 2, exon 4b, exon 4c or to the additional sequence thereof. Methods to derive specific primers from known sequences are well known in the art.

20 In another embodiment, the present invention relates to a kit for detecting the presence of differentially expressed PCA3 mRNA in a sample comprising at least one container means having disposed therein the above-described nucleic acid probe. In a preferred embodiment, the kit
25 further comprises other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound nucleic acid probe. Examples of detection reagents include, but are not limited to radiolabelled probes, enzymatic labeled probes (horse radish

peroxidase, alkaline phosphatase), and affinity labeled probes (biotin, avidin, or streptavidin).

In more detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the probe or primers used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, and the like), and containers which contain the reagents used to detect the hybridized probe, bound antibody, amplified product, or the like.

One skilled in the art will readily recognize that the nucleic acid probes described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art

In another embodiment of the present invention (and similarly to probes of the present invention) the antibodies of the present invention can be immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art. The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays as well as in immunochromatography.

In another embodiment, the present invention relates to a method of detecting a differentially expressed PCA3 polypeptide in a sample,

comprising: a) contacting the sample with an above-described antibody (or protein), under conditions such that immunocomplexes form, and b) detecting the presence of the antibody bound to the polypeptide. In detail, the methods comprise incubating a test sample with one or more of the antibodies of the present invention and assaying whether the antibody binds to the test sample. The relative levels of differentially expressed PCA3 in a sample enable a distinction between a malignant and non-malignant prostatic state.

In a further embodiment, the present invention relates to a method of detecting a PCA3 antibody in a sample, comprising a) contacting the sample with an above-described differentially expressed PCA3 protein, under conditions such that immunocomplexes form, and b) detecting the presence of the protein bound to the antibody or antibody bound to the protein. In detail, the methods comprise incubating a test sample with one or more of the proteins of the present invention and assaying whether the antibody binds to the test sample.

In another embodiment of the present invention, a kit is provided which contains all the necessary reagents to carry out the previously described methods of detection.

The kit can comprise i) a first container means containing an above-described antibody, and ii) second container means containing a conjugate comprising a binding partner of the antibody and a label

The kit can comprise: i) a first container means containing an above-described protein, and preferably, ii) second container means containing a conjugate comprising a binding partner of the protein and a label. More specifically, a diagnostic kit comprises a differentially expressed PCA3 protein as described above, to detect antibodies in the serum of potentially infected animals or humans

In another preferred embodiment, the kit further comprises one or more other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound antibodies. Examples of detection reagents include, but are not limited to, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the chromophoric, enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. The compartmentalized kit can be as described above for nucleic acid probe kits.

One skilled in the art will readily recognize that the antibodies described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

It is to be understood that although the following discussion is specifically directed to human patients, the teachings are also applicable to any animal that differentially expresses PCA3 mRNAs.

The diagnostic and screening methods of the invention are especially useful for a patient suspected of being at risk for developing a disease associated with an altered expression level of PCA3 based on family history, or a patient in which it is desired to diagnose a PCA3-related disease (ex. prostate cancer).

According to the invention, presymptomatic screening of an individual in need of such screening is now possible using DNA encoding the PCA3 protein or the PCA3 gene of the invention or fragments thereof. The screening method of the invention allows a presymptomatic diagnosis, of the presence of the PCA3-minus additional sequence, differentially expressed PCA3 mRNA in individuals, and thus an opinion concerning the likelihood that such individuals would develop or have developed a PCA3-associated disease or have a normal prostatic state. This is especially valuable for the identification of carriers of altered PCA3 genes, for example, from individuals

with a family history of a PCA3-associated disease. Early diagnosis is also desired to maximize appropriate timely intervention.

In one preferred embodiment of the method of screening, a tissue sample would be taken from such individual, and screened for (1) the presence of the PCA3 mRNA lacking the additional sequence of the present invention; (2) the presence of the additional sequence-containing PCA3 mRNA and/or (3) the presence of differentially expressed PCA3 protein. PCA3 mRNA can be characterized and compared to determine differentially expressed PCA3 mRNA (a) levels and/or (b) size. Lastly, differentially expressed PCA3 protein can be (a) detected and/or (b) quantitated using a biological assay for PCA3 activity or using an immunological assay and PCA3 antibodies. A presence of a PCA3 mRNA lacking the additional sequence (or an mRNA not modifying and/or not interrupting the PCA3 coding sequence) and/or of the protein encoded thereby would indicate that the patient is at risk for developing prostate cancer, or has developed prostate cancer. A presence of a PCA3 mRNA containing the additional sequence of the present invention and/or of the protein encoded thereby, in the absence of PCA3 mRNA lacking the additional sequence and/or the protein encoded thereby or at a level superior to that of the mRNA lacking the additional sequence and/or the protein encoded thereby would indicate that the patient has not yet developed prostate cancer, and/or has a lower risk of developing prostate cancer.

Therapeutic effects of therapeutic nucleic acids can include, but are not limited to turning off or modifying the processing of the differentially expressed PCA3 mRNA lacking the additional sequence of the present invention. In addition, an expression of a differentially-expressed PCA3 mRNA comprising the additional sequence in accordance with the

present invention to a higher level than that of the PCA3mRNA lacking the additional sequence could have cancer-reversing effects on cells.

Included as well in the invention are pharmaceutical compositions comprising an effective amount of at least one antisense oligonucleotide to a PCA3 mRNA lacking the additional sequence, in combination with a pharmaceutically acceptable carrier. Such antisense oligos include, but are not limited to, at least one nucleotide sequence of 12-500 bases in length which is complementary to at least a portion of SEQ ID NO:2.

Thus, broadly, the invention provides means to shift the balance between the quantity of the differentially expressed PCA3mRNAs such that the malignant state of a cell can be modulated.

Specificity for gene expression in prostate cancer cells can be conferred by using appropriate cell-specific regulatory sequences, such as cell-specific enhancers and promoters. Thus, gene therapy can be used to alleviate PCA3 related pathology by inhibiting the inappropriate expression of a particular form of PCA3. Moreover, gene therapy can be used to alleviate such pathologies by providing the appropriate expression level of a particular form of PCA3. In this case, particular PCA3 nucleic acid sequences can be coded by DNA or RNA constructs which are administered in the form of viruses, as described above.

The present invention provides the above-described PCA3 antibodies (preferably, PCA3 murine antibodies and chimeric PCA3 murine-human antibodies, and fragments and regions thereof) which inhibit or neutralize PCA3 biological activity *in vivo* and are specific for PCA3. These antibodies can be used for therapeutic purposes in subjects having pathologies or conditions associated with the presence of aberrant PCA3 expression. Antibodies, and fragments, regions and derivatives thereof, of

the present invention preferably contain at least one region which recognizes an epitope of PCA3 which has inhibiting and/or neutralizing biological activity *in vivo*.

5 Treatment comprises parenterally administering a single or multiple doses of the antibody, fragment or derivative Preferred for human pharmaceutical use are high affinity potent PCA3-inhibiting and/or neutralizing murine and chimeric antibodies, fragments and regions of this invention.

10 Monoclonal antibodies of the present invention may be administered by any means that enables the active agent to reach the agent's site of action in the body of a mammal. Because proteins are subject to being digested when administered orally, parenteral administration, i.e., intravenous, subcutaneous, intramuscular, would ordinarily be used to optimize absorption.

15 Monoclonal antibodies of the present invention may be administered either as individual therapeutic agents or in combination with other therapeutic agents They can be administered alone, but are generally administered with a pharmaceutical carrier selected on the basis of the chosen route of administration and standard pharmaceutical practice.

20 The dosage administered will, of course, vary depending upon known factors such as the pharmacodynamic characteristics of the particular agent, and its mode and route of administration; age, health, and weight of the recipient; nature and extent of symptoms, kind of concurrent treatment, frequency of treatment, and the effect desired. Usually a daily
25 dosage of active ingredient can be about 0.1 to 100 milligrams per kilogram of body weight. Ordinarily 0.5 to 50, and preferably 1 to 10 milligrams per kilogram per day given in divided doses 1 to 6 times a day or in sustained release form is effective to obtain desired results.

Dosage forms (composition) suitable for internal administration generally contain from about 1 milligram to about 500 milligrams of active ingredient per unit. In these pharmaceutical compositions the active ingredient will ordinarily be present in an amount of about 0.5-95% by weight based on the total weight of the composition.

Cytotoxic drugs which can be conjugated to antibodies and subsequently used for *in vivo* therapy include, but are not limited to, daunorubicin, doxorubicin, methotrexate, and Mitomycin C.

The non-human animals of the invention comprise any animal having a transgenic interruption or alteration of the endogenous gene(s) (knock-out animals) and/or into the genome of which has been introduced one or more transgenes that direct the expression of differentially expressed human PCA3 mRNAs. Also preferred are the introduction of antisense PCA3 nucleic acids.

Such non-human animals include vertebrates such as rodents, non-human primates, sheep, dog, cow, amphibians, reptiles, etc.

Preferred non-human animals are selected from non-human mammalian species of animals, most preferably, animals from the rodent family including rats and mice, most preferably mice.

The transgenic animals of the invention are animals into which has been introduced by nonnatural means (i.e., by human manipulation), one or more genes that do not occur naturally in the animal, e.g., foreign genes, genetically engineered endogenous genes, etc. The nonnaturally introduced genes, known as transgenes, may be from the same or a different species as the animal but not naturally found in the animal in the configuration and/or at the chromosomal locus conferred by the transgene. Transgenes may comprise foreign DNA sequences, i.e., sequences not normally found in the genome of the host animal. Alternatively

or additionally, transgenes may comprise endogenous DNA sequences that are abnormal in that they have been rearranged or mutated *in vitro* in order to alter the normal *in vivo* pattern of expression of the gene, or to alter or eliminate the biological activity of an endogenous gene product encoded by the gene.

The transgenic non-human animals of the invention are produced by introducing transgenes into the germline of the non-human animal. Embryonic target cells at various developmental stages are used to introduce the transgenes of the invention. Different methods are used depending on the stage of development of the embryonic target cell(s). These methods are well known in the art.

Transgenes may be introduced into non-human animals in order to provide animal models for human diseases. Transgenes that result in such animal models include, e.g., transgenes that encode the differentially expressed PCA3 mRNAs associated with a malignant prostate status (i.e., prostate cancer) or a non-malignant prostate status.

Having identified a marker sequence in differentially expressed PCA3 mRNA and a correlation between the balance of the expression level of the differentially expressed PCA3 mRNAs (or protein encoded thereby) and the malignant or non-malignant prostatic states, the present invention opens the way to numerous methods, assays and reagents for the prognosis, diagnosis, staging, predisposition and therapy of prostate cancer. In a broad embodiment, the present invention provides the means to assess prostate cancer by identifying PCA3 mRNA lacking the additional sequence in accordance with the present invention (or a protein encoded thereby). Numerous methods, primers, probes, antibodies and reagents can be used to identify such a nucleic acid molecule (or such a protein), as will be clear to the skilled artisan to which the present invention pertains.

The present invention is illustrated in further detail by the following non-limiting examples.

EXAMPLE 1

Identification of differentially expressed PCA3 mRNAs and correlation of their expression with prostatic disease

PCA3-specific PCR primers were developed in order to analyze PCA3 expression in different samples. To be able to discern between sequences amplified from mRNA (messenger RNA) and genomic DNA, these primers were designed to span an intron, in occurrence intron 3. As illustrated in Figure 1, the PCA3 sense primer lies within exon 3 and the PCA3 antisense primer within exon 4a. Samples to be analyzed for PCA3 expression consisted of frozen tissue chips removed by transurethral resection of the prostate (BPH, 4 patients) or frozen prostates obtained by radical prostatectomy (prostate cancer, 6 patients). Radical prostatectomy samples were processed into frozen sections to specifically select regions containing prostate cancer cells. RNA was extracted from the frozen samples using a liquid-phase RNA extraction method (Trizol). Extracted nucleic acids were subsequently treated with DNase in order to digest genomic DNA. DNase-treated RNA was reverse transcribed into cDNA using reverse transcriptase and then submitted to PCR-analysis using the PCA3 primers. PCR was performed for 35 cycles with Taq DNA polymerase, amplified material was separated on agarose gels and visualized by ethidium bromide staining. As shown in Figure 2, PCR amplification of PCA3 generates two products which can be separated by size and differ in relative abundance. The smaller amplicon (277 bp) is predominantly or exclusively found in samples from prostate cancer patients (Figure 2, upper row) whereas the larger amplicon (505 bp) is more prominent in samples from patients with a

non-malignant prostatic state (BPH [Figure 2, lower row]). Pathological examination of patients' biopsies confirmed the initial diagnosis for each patient except for Patient BPH 1 which was found to have prostate cancer.

In order to confirm the origin of the amplified fragments, they were isolated from the gel and sequenced. Sequences are shown in Figure 3. As expected, the smaller 277 bp fragment proved to correspond to the regions of exons 3 and 4a spanned by the PCA3 PCR primers. The larger 505 bp fragment is identical to the smaller fragment except for the herein identified sequence which lies between exon 3 and exon 4a. Of note, direct PCR analysis of all samples without reverse transcription did not yield amplified material ruling out the hypothesis that the larger amplification product originates from genomic DNA

Thus, PCA3 mRNA is present in at least two distinct forms within the cell, a short form lacking the herein identified additional sequence (hereafter called sequence 2; SEQ ID NO:2) as well as a long form having this additional sequence (hereafter called sequence 1; SEQ ID NO:1). The presence of the additional sequence in sequence 1 interrupts the predicted open reading frame coding for the PCA3 protein. The predicted sequence of the protein encoded by this long PCA3 mRNA is shown in Figure 4. As illustrated in Figure 2, relative expression levels of the two PCA3 mRNA sequences vary dependent on the cell type. Prostate cancer cells predominantly express sequence 2 whereas BPH cells mainly express sequence 1.

These observations demonstrate that it is possible to discern between a malignant and non-malignant state of a prostate. As well, it is tempting to predict that the relative levels of the two types of PCA3 mRNAs will enable to discern the benign state from the malignant state.

EXAMPLE 2**Assessment of the prostatic state of
a patient using RT-PCR**

5 Patient samples were obtained and RNA prepared therefrom
as commonly known. Reverse transcription mixes were prepared as RT
follows: 0.2 μ g total RNA + 0.6 μ g pdN6 (random hexamer primers) + 1.25
mM dNTPs + 200 U M-MLV reverse transcriptase in 50 mM Tris-HCl pH 8.3,
75 mM KCl, 3 mM MgCl₂, 10 mM DTT. The mixture was incubated 1 hr at
40°C.

10 4 μ l of the RT-reaction of above was mixed in 50 μ L of 20 mM
Tris-HCl pH 8.4, 50 mM KCl, 2.5 mM MgCl₂, 0.5 mM dNTPs, 0.5 μ M of each
primer and 2.5 U Taq DNA polymerase. For PCR analysis, the amplification
was carried out for 35 cycles (1 min each at 94°C, 60°C, 72°C) followed by
a 10 min extension at 72°C. The PCR products were analyzed by
15 conventional agarose gel electrophoresis.

Although the present invention has been described
hereinabove by way of preferred embodiments thereof, it can be modified,
without departing from the spirit and nature of the subject invention as
defined in the appended claims.